

A SIMPLE MODEL OF A LIMIT ORDER BOOKELENA YUDOVINA,^{*} *University of Michigan***Abstract**

We formulate a simplified model of a limit order book, in which the arrival process is independent of the current state. We prove a phase transition result: there exist prices κ_b and κ_a such that, for any $\epsilon > 0$, only finitely many bid (ask) departures occur at prices below $\kappa_b - \epsilon$ (above $\kappa_a + \epsilon$), while the interval $(\kappa_b + \epsilon, \kappa_a - \epsilon)$ infinitely often contains no bids, and infinitely often contains no asks. We derive expressions for κ_b and κ_a , which we solve in the case of uniform arrivals. We conjecture the positive recurrence of a modified model, and find the steady-state distribution of the highest bid and of the lowest ask assuming the positive recurrence.

Keywords: limit order book, Lyapunov function, limiting distribution

2010 Mathematics Subject Classification: Primary 60J20, 91B26

Secondary 60K25

1. Introduction

A limit order book is a pricing mechanism for a single-commodity market, in which users can trade off time against price by submitting orders to be executed at a later time, once the price becomes acceptable. This mechanism is used in many financial markets, and has generated extensive research, both empirical and theoretical. We do not aim to give an overview of the field here; references can be found in the survey by Gould et al. [5].

While much of the research has been either empirical studies of real-world markets, or game-theoretic analysis, our approach is to consider a Markovian model. This avoids

^{*} Postal address: Department of Statistics

439 West Hall

1085 South University Ave.

Ann Arbor, MI 48109

Email address: yudovina@umich.edu

the difficulties of prescribing models of individual user behaviour by assuming certain stochastic dynamics for the market as a whole. The pioneering paper of Gode and Sunder [4] showed that many of the features of a market may be reproduced even with zero-intelligence traders. Our model is somewhat similar to the models considered by Cont and de Larrard [2], Cont et al. [3], and Simatos [6]; however, the set-up differs from their work because we model the arrival events as independent of the state of the system. This assumption can be interpreted as treating the system on relatively short time scales, where the price does not significantly change. We discuss this at greater length in Remark 2, after formulating our model.

It is surprising that even in such a simple setting, nontrivial behaviour emerges. Specifically, we find that the system experiences a phase transition: at prices below a certain threshold, only finitely many bid orders will ever be executed; at prices above the threshold, all the queued orders will clear infinitely often. (And similarly for ask orders, of course.) The probabilistic techniques used in this paper are not difficult, and showcase the fact that our model is attractive and amenable to analysis; we outline some of the extensions that could be considered in Section 9.

1.1. Basic notation

For a process ξ_t indexed by time, ξ_{t-} indicates the state of ξ just before t . We will take all our processes to be right-continuous.

For a set A , $\mathbf{1}_A$ is the indicator function of A . For two sets A and B , $A \Delta B$ is the symmetric difference $(A \setminus B) \cup (B \setminus A)$.

2. Model

The basic dynamics of the system are as follows. At time 0, the limit order book is empty. Limit bids and asks arrivals form two independent point processes in $\mathbb{R} \times \mathbb{R}_+$. Arriving orders are iid; in particular, the interarrival times, types, and prices of the arriving orders are independent of each other and of the state of the limit order book. For convenience, we will assume that the distributions of prices of arriving orders are absolutely continuous, and write f_b , respectively f_a , for the density of the price of the arriving bid, respectively ask. We write F_b and F_a for the cumulative density. For

the interarrival times, we assume that the event “infinitely many orders arrive; no two orders arrive at the same time; finitely many orders arrive over any finite time interval” has probability 1. In all that follows, we work on the probability-1 event that all order arrival times and prices are distinct.

The state of the limit order book at time t is the two counting measures of bid orders and ask orders present in the limit order book. (Under our assumption, this is simply the set of prices of bids and asks.) Additionally, we keep track of the highest bid price β_t and lowest ask price α_t inside the limit order book. We define $\beta_t = -\infty$ if there are no bids inside the book, and similarly $\alpha_t = \infty$ if there are no asks.

The change to the limit order book that occurs upon arrival depends on the location of the price of the arriving order relative to these two prices. If the arriving order at time t is an ask at price p , then:

(i) If $p < \beta_{t-}$, the newly-arrived ask causes the bid at price β_{t-} to be executed and leave. In this case, $\beta_t < \beta_{t-}$, and the ask side is unchanged. (ii) If $\beta_{t-} < p < \alpha_{t-}$, the newly arrived ask joins the limit order book, and $\alpha_t = p$. (iii) If $\alpha_{t-} < p$, the newly arrived ask joins the limit order book, and $\alpha_t = \alpha_{t-}$.

Similarly, if the arriving order at time t is a bid at price q , then:

(i) If $q > \alpha_{t-}$, the newly-arrived bid causes the ask at price α_{t-} to be executed and leave. In this case, $\alpha_t > \alpha_{t-}$, and the bid side is unchanged. (ii) If $\alpha_{t-} > q > \beta_{t-}$, the newly arrived bid joins the limit order book, and $\beta_t = q$. (iii) If $\beta_{t-} > q$, the newly arrived bid joins the limit order book, and $\beta_t = \beta_{t-}$.

Abandonments are not allowed. Thus, bids and asks may depart if they are executed, or remain in the limit order book forever.

It follows from these dynamics that $\beta_t < \alpha_t$ always, i.e. all the bid orders in the limit order book are to the left of all the ask orders.

Remark 1. A convenient way to interpret bids that arrive at prices above α_t is to think of all of them as *market bid* orders arriving at the current best price α_t , and similarly for asks arriving at prices below β_t . The rate at which market bid orders arrive will depend on the current lowest ask price α_t ; it will increase when α_t is low, and decrease when α_t is high (and similarly for asks). In particular, the total rate at which market orders arrive will be higher when the bid-ask spread $\alpha_t - \beta_t$ is low, and

lower when it is high.

Remark 2. This model differs from a real limit order book in several important aspects, which we now discuss.

First, ignoring abandonments represents a great difference to real-world markets, where a large fraction of the orders are canceled before being executed. However, if we consider the model only over relatively short time scales, then orders that are eventually canceled may be treated as “remaining in the system forever”, while orders that are canceled very quickly may be interpreted as background noise. Properly incorporating abandonments into our model would be difficult, because the results of Section ?? rely on only the best orders departing, and then doing so in bid-ask pairs.

Second, we consider the arrival process of the orders to be independent from the state of the limit order book. Allowing the arrivals to depend on the state of the book would require a better understanding of the book’s shape; this is work in progress. However, over relatively short time scales, during which the price does not shift substantially, we may expect this assumption to be reasonable.

Last, we consider the arriving orders to all have size 1. Allowing non-unit-sized orders also requires a better understanding of the dynamics of the shape of the limit order book, because a large arriving order may substantially move the highest price. In particular, in binned models (defined below), a large arriving order may remove orders from several bins, as opposed to just one. Our analysis can be extended without substantial change to accommodate orders of the form “buy n units or all of the orders available at the best market price, but do not buy any orders at higher prices”.

2.1. Modifications

We will consider a variant of the model with a finite number of price ticks, or bins. We partition \mathbb{R} into some number (possibly infinite) of disjoint convex nonempty subsets (i.e. points or intervals). We will consider two versions of binned models:

- (a) Ordinary binned limit order book: the arriving bid at price p is allowed to depart with the lowest ask if p and α_{t-} fall into the same bin (even if $p < \alpha_{t-}$), or if $p > \alpha_t$, and similarly for arriving asks.
- (a) Strict binned limit order book: the arriving bid at price p is allowed to depart with the lowest ask only if $p > \alpha_{t-}$ and they fall into different bins, and similarly for

arriving asks.

In the binned models, we can treat the arrival price distribution as being supported on the set of bins. However, for coupling arguments it will be convenient to think of arrivals as coming from an underlying continuous distribution on \mathbb{R} .

Additionally, we may consider non-zero initial states of the limit order book. In particular, we may allow the initial state to have an infinite number of bids or asks at a certain price, as long as always all bids are lower than all asks. This has the interpretation of a large player in the market, who is offering infinite liquidity at some (low) buy price, and some (high) ask price. Note that if we have infinitely many bids at some price p , then bids below p will never leave the system, while arriving asks below p will always immediately depart. Thus, we may focus our attention only on the prices above p .

Remark 3. (*Coordinate transformation.*) It will sometimes be convenient for us to change coordinates so that the bids and asks arrive on $[0, 1]$, and, moreover, the bid distribution f_b is uniform over $[0, 1]$. This can be done, e.g., by applying the transformation $x \mapsto F_b(x)$.

3. Results

We now state our main results.

Theorem 3.1. *For any of the variants of the limit order book discussed above, there exist deterministic constants κ_b and κ_a with the following properties. For any $\epsilon > 0$,*

- $\beta_t < \kappa_b - \epsilon$ occurs only finitely many times; $\beta_t < \kappa_b + \epsilon$ occurs infinitely often. Thus, bids below $\kappa_b - \epsilon$ eventually never leave, while above $\kappa_b + \epsilon$ infinitely often there are no bids.
- Similarly, $\alpha_t > \kappa_a + \epsilon$ occurs only finitely many times; $\alpha_t > \kappa_a - \epsilon$ occurs infinitely often. Thus, asks above $\kappa_a + \epsilon$ eventually never leave, while below $\kappa_a - \epsilon$ infinitely often there are no asks.

This indicates a sharp phase transition in the behaviour of the orders at low, medium, and high prices. We will identify the threshold values κ_b and κ_a below.

The following alternative characterization of κ_b and κ_a will be useful. For a limit order book \mathcal{L} , let $B_t(p)$ denote the number of bids at time t at prices $\leq p$, and let $A_t(p)$ denote the number of asks at time t at prices $\geq p$. Note that asks are counted from the right. Clearly, we have $B_t(\infty) = B_t(\beta_t)$ and similarly $A_t(-\infty) = A_t(\alpha_t)$.

Corollary 3.1. *Suppose the arrival process is Poisson of rate 1 in time, and the arrival price distributions are continuous. The values of κ_b and κ_a may be found as*

$$F_b(\kappa_b) = \liminf_{T \rightarrow \infty} \frac{1}{T} B_T(\infty)$$

and

$$1 - F_a(\kappa_a) = \liminf_{T \rightarrow \infty} \frac{1}{T} A_T(-\infty).$$

The proof of these results appears in Section 4.

The surprising fact is that we can obtain numeric values of κ_b and κ_a in terms of the arrival distributions.

Theorem 3.2. *Suppose arriving orders are equally likely to be bids and asks, and the densities f_b and f_a are absolutely continuous with respect to each other. Suppose further that κ_b and κ_a are known to be finite; for example, this is the case if there exist $x < y$ with the property that $F_b(x) = 1 - F_a(y)$, $F_b(y) = 1 - F_a(x)$, and $F_b(y) > 1 - 2F_b(x)(1 - F_b(x))$. (See Lemma 6.3)*

Then the threshold values κ_b and κ_a are the unique pair of finite numbers satisfying $F_b(\kappa_b) = 1 - F_a(\kappa_a)$, such that the solution ϖ^b of the second-order ODE

$$\left(-\frac{f_a(x)}{1 - F_b(x)} (F_a(x)\varpi^b(x))' \right)' = \varpi^b(x)f_b(x)$$

with initial conditions

$$\varpi^b(\kappa_b) = \frac{1}{F_a(\kappa_b)}, \quad \frac{d}{dx} \varpi^b(x)|_{x=\kappa_b} = -\frac{f_a(\kappa_b)}{F_a(\kappa_b)^2}$$

satisfies $\varpi^b(x) \rightarrow 0$ as $x \uparrow \kappa_a$.

If $f_b = f_a = \mathbf{1}_{[0,1]}$, then

$$\varpi^b(x) = (1 - \kappa) \left(\frac{1}{x} + \log \left(\frac{1 - x}{x} \right) \right), \quad x \in (\kappa, 1 - \kappa)$$

where if w is the unique solution to $we^w = e^{-1}$, then $\kappa = \frac{w}{w+1} \approx 0.217$.

The proof of this result is in Section 7.

We conjecture that $\varpi^b(x)f_b(x)$ is the steady-state density of the distribution of β_t (with respect to the Lebesgue measure). Unfortunately, we have been unable to show the positive recurrence that would imply the existence of a steady-state density for β_t , so instead in Lemma 6.2 and proof of Theorem 3.2 we derive that this quantity is the ergodic limit of the empirical distribution of β_t along a certain sequence of times. We conjecture that the true result is as follows.

Conjecture 1. *Let the arrival process be as in Theorem 3.2.*

1. *Consider a binned limit order book with infinitely many bids in the bin containing κ_b , and infinitely many asks in the bin containing κ_a . (Its state is described by the number of and type of orders in the bins between these two.) This limit order book is recurrent.*
2. *Let $\epsilon > 0$ be fixed. Consider a limit order book \mathcal{L} whose initial state has infinitely many bids at $\kappa_b + \epsilon$ and infinitely many asks at $\kappa_a - \epsilon$. (If κ_b and κ_a are infinite, put the bids and asks at $F_b^{-1}(\epsilon)$ and $F_a^{-1}(1 - \epsilon)$.) This limit order book is positive Harris recurrent.*

Our analysis of ϖ^i , $i = a, b$ in Lemma 6.2 and the proof of Theorem 3.2 will show the following corollary.

Corollary 3.2. *Suppose Conjecture 1 holds. Let ϖ_ϵ^b and ϖ_ϵ^a be the distribution of the rightmost bid and the leftmost ask in the limit order books with infinitely many bids at $\kappa_b + \epsilon$ and asks at $\kappa_a - \epsilon$. As $\epsilon \rightarrow 0$, we have $\varpi_\epsilon^b(x) \rightarrow \varpi^b(x)f_b(x)$ and $\varpi_\epsilon^a(x) \rightarrow \varpi^a(x)f_a(x)$ uniformly.*

Evidence (theoretical and numerical) supporting the conjecture is presented in Section 8.

4. Coupling and monotonicity

We now present some coupling arguments, which show monotonicity properties of our system. Our results will compare behaviors of limit order books \mathcal{L} and $\tilde{\mathcal{L}}$ with the same underlying arrival process; we will consider the effect of changing the initial state

and the effect of changing the bid-ask matching rule by changing the binning. We will refer to the state of the limit order books at time t as L_t and \tilde{L}_t respectively.

4.1. Initial state

Lemma 4.1. *Let \mathcal{L} and $\tilde{\mathcal{L}}$ be two limit order books with the same arrival process and order matching rule (i.e., both ordinary, or both binned with the same bins, or both strict binned with the same bins). Suppose the initial state \tilde{L}_0 differs from L_0 by the addition of a single bid. Then at all times t , the state \tilde{L}_t differs from L_t either by the addition of a single bid, or by the removal of a single ask. Similarly, if \tilde{L}_0 differs from L_0 by the addition of a single ask, then \tilde{L}_t differs from L_t either by the addition of a single ask or removal of a single bid.*

Proof. We prove the statement for the case of an extra bid, the case of an extra ask being entirely similar. The proof proceeds by induction on the number of arriving orders.

Clearly the statement is true before any orders arrive. Moreover, until the extra bid is removed in $\tilde{\mathcal{L}}$, the order arrivals and departures in \mathcal{L} and $\tilde{\mathcal{L}}$ coincide. Consider the time when the extra bid is removed in $\tilde{\mathcal{L}}$; this corresponds to the arrival of an ask at some price p . Now, if in \mathcal{L} this ask also immediately departs (with some other bid at price q), then the state of $\tilde{\mathcal{L}}$ differs from the state of \mathcal{L} by the addition of a bid (at price q). If, however, in \mathcal{L} the ask does not immediately depart, then the state of $\tilde{\mathcal{L}}$ differs from the state of \mathcal{L} by the removal of this ask.

We obtain some easy, but useful corollaries.

Corollary 4.1. *Consider a limit order book \mathcal{L} , and construct $\tilde{\mathcal{L}}$ by, at some finite number of points in time, adding or removing a finite number of orders from \mathcal{L} , for a total of at most M . Then at all times, the states of \mathcal{L} and $\tilde{\mathcal{L}}$ differ by at most M orders.*

Corollary 4.2. *Consider a limit order book \mathcal{L} , and construct $\tilde{\mathcal{L}}$ by, at some points in time, adding some number of bids (but leaving the asks unchanged). Then at all times $\tilde{\mathcal{L}}$ contains all of the bids in \mathcal{L} (and possibly some more), and a subset of the asks. Similarly, if we add some number of asks, but leave the bids unchanged, then $\tilde{\mathcal{L}}$ will contain all of the asks in \mathcal{L} , and a subset of the bids.*

We are now in a position to prove Theorem 3.1.

Proof of Theorem 3.1. Our goal is to show that the event

$$\mathcal{E}(x) = \{\beta_t < x \text{ infinitely often}\}$$

occurs with probability 0 or 1 for any x . Note that for $x < y$ we have $\mathcal{E}(x) \subset \mathcal{E}(y)$; we will take $\kappa_b = \inf\{x : \mathbb{P}(\mathcal{E}(x)) = 1\}$.

For $M \geq 0$, let

$$\mathcal{E}^M(x) = \{\beta_t < x \text{ infinitely often in } \tilde{\mathcal{L}}\},$$

where $\tilde{\mathcal{L}}$ is the limit order book whose initial state is the same as that of \mathcal{L} , but the first M arrivals do not happen. We will show that $\mathcal{E}(x) = \mathcal{E}^M(x)$ with probability 1. Thus, $\bigcap_M \mathcal{E}^M(x)$ is a tail event which coincides with $\mathcal{E}(x)$ with probability 1. By Kolmogorov's 0-1 law, $\mathbb{P}(\bigcap_M \mathcal{E}^M(x)) \in \{0, 1\}$, which proves the result.

We now show $\mathcal{E}(x) = \mathcal{E}^M(x)$ almost surely. By Corollary 4.1, along every trajectory the states of \mathcal{L} and $\tilde{\mathcal{L}}$ differ by at most M orders. In particular,

$$\beta_t \leq x \implies \tilde{B}_t(x) \leq M,$$

and conversely, $\tilde{\beta}_t \leq x \implies B_t(x) \leq M$. (Recall $B_t(x)$ counts the number of bids at time t at prices $\leq x$.)

Clearly, for $x < \text{ess inf}(f_a)$, neither \mathcal{E} nor \mathcal{E}^M occur, since no bid departures can happen. Therefore, let $x > \text{ess inf}(f_a)$. Whenever $B_t(x) \leq M$, the conditional probability that M order arrivals later we will have

$$\beta_{(t+M \text{ interarrival times})} < x$$

is bounded below uniformly in t ; and similarly for \tilde{B} and $\tilde{\beta}$. Consequently, $\mathbb{P}(\mathcal{E}(x) \Delta \mathcal{E}^M(x)) = 0$ as required.

The proof for κ_a is entirely similar.

We now prove Corollary 3.1.

Proof of Corollary 3.1. Pick $\epsilon > 0$. By the definition of κ_b , and the strong law of large numbers for the arrival process, we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} B_T(\kappa_b - \epsilon) = F_b(\kappa_b - \epsilon).$$

Since the number of bids elsewhere in the book is nonnegative, we obtain

$$\liminf_{T \rightarrow \infty} \frac{1}{T} B_T(\infty) \geq F_b(\kappa_b - \epsilon).$$

Moreover, we know that there exists a sequence of times $T_n \rightarrow \infty$ along which $\beta_{T_n} < \kappa_b + \epsilon$. Consequently,

$$\liminf_{T \rightarrow \infty} \frac{1}{T} B_T(\infty) \leq F_b(\kappa_b + \epsilon).$$

Since ϵ is arbitrary and we assumed that F_b is continuous, this proves the result. The result for κ_a is proved entirely similarly.

Note that we only needed F_b to be continuous at κ_b and F_a to be continuous at κ_a .

4.2. Binning

Before presenting the formal results in this section, we give the intuition. Consider a binned limit order book; recall that in an (ordinary) limit order book, a bid-ask pair may leave if they are in the same bin, even if the bid price is lower than the ask price. Now suppose we make bins larger. Intuitively, this should make it easier for bid-ask pairs to leave, so we expect to find fewer bids and asks in the system. The intuition is reversed for limit order books, where bid-ask pairs in the same bin are *not* allowed to leave: there, making bins larger should leave more unfulfilled orders.

For two different partitions $\Pi, \tilde{\Pi}$ of \mathbb{R} into bins, we say that Π refines $\tilde{\Pi}$ if every bin of $\tilde{\Pi}$ is the union of one or more bins of Π .

Lemma 4.2. *Let \mathcal{L} and $\tilde{\mathcal{L}}$ be two ordinary binned limit order books with the same initial state and arrival process, and suppose that the binning partition Π of \mathcal{L} refines the partition $\tilde{\Pi}$ of $\tilde{\mathcal{L}}$. Then at all times t and prices p we have*

$$\tilde{B}_t(p) \leq B_t(p), \quad \tilde{A}_t(p) \leq A_t(p).$$

Proof. The proof proceeds again by induction on the number of arrived orders. We show the inequality for $\tilde{B}_t \leq B_t$. Clearly it holds before any orders arrive.

Suppose at time t , an arrival of a bid at price p occurs. In order to destroy the inequality $\tilde{B} \leq B$, we would need to have p join the book in $\tilde{\mathcal{L}}$ but leave in \mathcal{L} , i.e. we must have

$$\alpha_{t-} \lesssim p \prec \tilde{\alpha}_{t-}.$$

Here, we mean that either $\alpha_{t-} < p$ or they occur in the same Π -bin, and $\tilde{\alpha}_{t-} > p$ and occurs in a different $\tilde{\Pi}$ -bin. Since Π is finer than $\tilde{\Pi}$, we must have $\tilde{\alpha}_{t-} > \alpha_{t-}$.

Furthermore, in order to destroy $\tilde{B} \leq B$ with the arrival of a single bid at price p , we must have had $\tilde{B}_{t-}(p) = B_{t-}(p)$ with equality. Note, however, that $B_{t-}(p) = B_{t-}(\infty)$, since the lowest ask in \mathcal{L} is in the same bin as p . Since at time $t-$ the inequality $\tilde{B}_{t-} \leq B_{t-}$ held, we have $\tilde{B}_{t-}(\infty) = B_{t-}(\infty)$.

Because bid-ask departures always occur in pairs, and the arrival processes were the same in $\tilde{\mathcal{L}}$ and \mathcal{L} , equality between total number of remaining bids implies equality for asks: $\tilde{A}_{t-}(-\infty) = A_{t-}(-\infty)$. Together with $\tilde{A}_{t-} \leq A_{t-}$, this implies $\tilde{\alpha}_{t-} \leq \alpha_{t-}$, and we've reached a contradiction. Thus, bid arrivals cannot destroy the inequality $\tilde{B} \leq B$.

Next, suppose that at time t an arrival of an ask at price p occurs. If this is to destroy $\tilde{B} \leq B$, then we must have

$$\tilde{\beta}_{t-} \prec p \lesssim \beta_{t-},$$

i.e. the ask leaves with the bid at β_{t-} in \mathcal{L} but does not remove a bid in $\tilde{\mathcal{L}}$. As before, these inequalities imply $\tilde{\beta}_{t-} < \beta_{t-}$.

Moreover, in order for the removal of the bid at β_{t-} to destroy the inequality $\tilde{B} \leq B$, we must have $\tilde{B}_{t-}(\beta_{t-}) = B_{t-}(\beta_{t-})$. Now, by definition $B_{t-}(\beta_{t-}) = B_{t-}(\infty)$, and since $\tilde{\beta}_{t-} < \beta_{t-}$, also $\tilde{B}_{t-}(\beta_{t-}) = \tilde{B}_{t-}(\infty)$. We conclude $\tilde{B}_{t-}(\infty) = B_{t-}(\infty)$, which (since $\tilde{B}_{t-} \leq B_{t-}$) implies $\tilde{\beta}_{t-} \geq \beta_{t-}$, a contradiction. Thus, ask arrivals also cannot destroy the inequality $\tilde{B} \leq B$, and we are done.

Entirely similarly, we can prove the corresponding statement for strict binned order books, in which the inequalities are reversed. We record the statement here for future reference.

Lemma 4.3. *Let \mathcal{L} and $\tilde{\mathcal{L}}$ be two strict binned limit order books with the same initial state and arrival process, and suppose that the binning partition Π of \mathcal{L} refines the partition $\tilde{\Pi}$ of $\tilde{\mathcal{L}}$. Then at all times t and prices p we have*

$$\tilde{B}_t(p) \geq B_t(p), \quad \tilde{A}_t(p) \geq A_t(p).$$

We obtain the following easy corollary.

Corollary 4.3. *Consider three binning partitions $\Pi, \tilde{\Pi}, \hat{\Pi}$ where Π refines both $\tilde{\Pi}$ and $\hat{\Pi}$. Let \mathcal{L} and $\tilde{\mathcal{L}}$ be ordinary binned limit order books with bin partitions Π and $\tilde{\Pi}$, and let $\hat{\mathcal{L}}$ be a strict limit order book with bin partition $\hat{\Pi}$. Let the initial states and arrival processes be the same, and let $\kappa_i, i = a, b$ be defined for the books as in Theorem 3.1. Then*

$$\hat{\kappa}_b \geq \kappa_b \geq \tilde{\kappa}_b, \quad 1 - \hat{\kappa}_a \geq 1 - \kappa_a \geq 1 - \tilde{\kappa}_a.$$

In the next section, we see that if the binning partitions are sufficiently fine, then $\hat{\kappa}_b$ and $\tilde{\kappa}_b$ are close to each other, which will allow us to compute the value of κ_b for an ordinary unbinned limit order book using finer and finer binning partitions.

5. Many-bin limit

In this section, our goal is to show that we can reduce the analysis of the ordinary limit order book to the analysis of binned models. This is easier, because we are then reduced to a countable state-space Markov chain.

We begin with a bound on the effect of changing the arrival process on the value of κ_b .

Lemma 5.1. *Consider two limit order books \mathcal{L} and $\tilde{\mathcal{L}}$ with the same matching rule, but different arrival processes. Let $p_i, f_i, i = a, b$ denote the probability that an arriving order in \mathcal{L} is of type i (bid or ask), and the density of the price of arriving orders. Let $\tilde{p}_i, \tilde{f}_i, i = a, b$ denote the corresponding quantities for $\tilde{\mathcal{L}}$. Finally, let $\kappa_i, \tilde{\kappa}_i, i = a, b$ be given as in Theorem 3.1. Then*

$$|\kappa_b - \tilde{\kappa}_b| \leq |p_b - \tilde{p}_b| \left(\|f_b - \tilde{f}_b\|_{TV} + \|f_a - \tilde{f}_a\|_{TV} \right),$$

and similarly for κ_a .

Proof. We will use the characterization of $\kappa_b = \liminf_{T \rightarrow \infty} \frac{1}{T} B_T(\infty)$ given in Corollary 3.1, together with Corollary 4.1. We will set up the maximal coupling between the arrival processes for \mathcal{L} and $\tilde{\mathcal{L}}$. First, we reparametrize time so that the arrival processes are Poisson, rate 1 in time, with densities $p_i f_i(p) dp \times dt$ for \mathcal{L} , and $\tilde{p}_i \tilde{f}_i dp \times dt$ for $\tilde{\mathcal{L}}$. Next, we construct the arrivals for \mathcal{L} and $\tilde{\mathcal{L}}$ using six independent Poisson processes

$P_b^{\text{com}}, P_b^1, P_b^2, P_a^{\text{com}}, P_a^1, P_a^2$ with densities

$$\begin{aligned} P_b^{\text{com}} : & (p_b \vee \tilde{p}_b)(f_b(p) \vee \tilde{f}_b(p))dp \times dt, & P_b^{\text{com}} : & (p_a \vee \tilde{p}_a)(f_a(p) \vee \tilde{f}_a(p))dp \times dt \\ P_b^1 : & (p_b - \tilde{p}_b)_+(f_b(p) - \tilde{f}_b(p))_+ dp \times dt, & P_a^1 : & (p_a - \tilde{p}_a)_+(f_a(p) - \tilde{f}_a(p))_+ dp \times dt \\ P_b^2 : & (\tilde{p}_b - p_b)_+(\tilde{f}_b(p) - f_b(p))_+ dp \times dt, & P_a^2 : & (\tilde{p}_a - p_a)_+(\tilde{f}_a(p) - f_a(p))_+ dp \times dt. \end{aligned}$$

Now, let the arrival processes for \mathcal{L} be given by $P_i^{\text{com}} + P_i^1$, $i = a, b$, and let the arrival processes for $\tilde{\mathcal{L}}$ be given by $P_i^{\text{com}} + P_i^2$. Note that the orders that arrive in one order book but not the other are simply the sum $P_i^1 + P_i^2$. Aggregating these orders over time, the total difference in arrivals constitutes a Poisson process whose rate is bounded above by

$$r = |p_b - \tilde{p}_b| \left(\|f_b - \tilde{f}_b\|_{TV} + \|f_a - \tilde{f}_a\|_{TV} \right).$$

Applying Corollary 4.1 together with the law of large numbers for Poisson processes, we see that this can change $\lim_{T \rightarrow \infty} \frac{1}{T} B_T(\infty)$ by at most r , hence the result.

This observation will allow us to compute the threshold values κ_i , $i = a, b$ for the ordinary limit order book using ordinary binned limit order books. Note that we already know how to bound the threshold values between ordinary and strict limit order books by Corollary 4.3. We now make the following observation.

Consider a strict binned order book with N bins. Consider also an ordinary limit order book with $N + 1$ bins, where bids arrive only in the leftmost N bins according to the bid arrival process of the strict order book, and asks arrive only in the rightmost N bins according to the ask arrival process of the strict order book. If we identify the leftmost N bins with the strict limit order book, then the cumulative bid counts $B_t(p)$ will coincide in the two models; if we identify the rightmost N bins instead, the cumulative ask counts $A_t(p)$ will coincide. If the partition into bins was quite fine, we expect these one-bin alterations should not make a large difference.

We now formalize this intuition.

Lemma 5.2. *Let \mathcal{L}^N be a sequence of ordinary, and $\tilde{\mathcal{L}}^N$ a sequence of strict binned limit order books. Assume that the arrival process is the same for all of the limit order books, and the distribution of prices of arriving bids and asks is continuous and supported on $[0, 1]$. Let the binning partition for $\tilde{\mathcal{L}}^N$ have the following properties: (i)*

none of the bins receive more than $1/N$ of all of the bids, (ii) none of the bins receive more than $1/N$ of all of the asks; (iii) none of the bins have width more than $1/N$. (E.g., place boundaries between bins at $\frac{i}{N}$, $F_b^{-1}(\frac{i}{N})$ and $F_a^{-1}(\frac{i}{N})$ for $i = 1, \dots, N$.) Let the binning partition for \mathcal{L}^N be the same as the binning partition of $\tilde{\mathcal{L}}^{N+1}$.

Define the sequences κ_i^N and $\tilde{\kappa}_i^N$, $i = a, b$ as in Theorem 3.1. Then, as $N \rightarrow \infty$,

$$|\kappa_b^N - \tilde{\kappa}_b^N| \rightarrow 0, \quad |\kappa_a^N - \tilde{\kappa}_a^N| \rightarrow 0.$$

Proof. By the discussion above the statement of the lemma, we know that the value of κ_b is the same in a strict limit order book with N bins and a non-strict limit order book with $N + 1$ bins and a slightly modified arrival process: we ignore at most $1/N$ of each of the bid and ask arrivals, and shift the remaining arrivals by one bin. To finish the proof, it remains to note that the total variation distance between the original and modified arrival distributions converges to 0, and the probability that an arriving offer is a bid in the modified arrival process converges to the original probability.

Of course, the same observations apply to other features of the limit order book under a similar scaling.

6. A limit along a subsequence

We now turn to examining the long-term distribution of the locations β_t and α_t . In all of the analysis in this section, we parametrize time so that bids and asks both arrive as Poisson processes of rate 1 in time; in particular, we assume that arriving orders are equally likely to be bids and asks.

We first prove the following slightly stronger description of κ_b and κ_a .

Lemma 6.1. *Almost surely there exists a sequence of times $T_n \rightarrow \infty$ along which*

1. $\frac{1}{T_n} B_{T_n}(\kappa_b) \rightarrow F_b(\kappa_b)$, $\frac{1}{T_n} A_{T_n}(\kappa_a) \rightarrow 1 - F_a(\kappa_a)$;
2. $\frac{1}{T_n} B_{T_n}(\infty) \rightarrow F_b(\kappa_b)$, $\frac{1}{T_n} A_{T_n}(-\infty) \rightarrow 1 - F_a(\kappa_a)$.

Proof. Recall that for any $\epsilon > 0$ there exists a sequence of times $T_n^\epsilon \rightarrow \infty$ along which $\beta_t < \kappa_b + \epsilon$, and hence $B_{T_n^\epsilon}(\infty) = B_{T_n^\epsilon}(\kappa_b + \epsilon)$. Since each of $\beta_t < \kappa_b - \epsilon$ and $\alpha_t > \kappa_a + \epsilon$ occur only finitely many times, we may without loss of generality assume

that these events do not occur after T_1^ϵ . Consequently, after T_1^ϵ , all bids arriving below $\kappa_b - \epsilon$ remain, and all asks arriving above $\kappa_b + \epsilon$ remain.

We now apply the law of large numbers to the bid and ask arrivals. Note that $B_t(x) - B_t(y)$ is always bounded above by the total number of bid arrivals between x and y , and similarly for asks. We conclude that, by shifting indices on the sequence T_n^ϵ , we can arrange the following:

1. $B_{T_n^\epsilon}(\infty) - B_{T_n^\epsilon}(\kappa_b) \leq (F_b(\kappa_b + \epsilon) - F_b(\kappa_b) + \epsilon)T_n^\epsilon$;
2. $(F_b(\kappa_b - \epsilon) - \epsilon)T_n^\epsilon \leq B_{T_n^\epsilon}(\kappa_b) \leq (F_b(\kappa_b) + \epsilon)T_n^\epsilon$;
3. $(1 - F_a(\kappa_a + \epsilon) - \epsilon)T_n^\epsilon \leq A_{T_n^\epsilon}(\kappa_a) \leq (1 - F_a(\kappa_a) + \epsilon)T_n^\epsilon$. Moreover, the total numbers of bids and asks in the system differs by $o(T_n^\epsilon)$, and in fact, $O(\sqrt{T_n^\epsilon})$. Indeed, the difference in the number of arrivals of bids and asks is clearly the magnitude of a symmetric random walk; and they always depart in pairs. Recalling that $F_b(\kappa_b) + F_a(\kappa_a) = 1$, the above imply (decreasing ϵ if necessary) the fourth condition,
4. $A_{T_n^\epsilon}(-\infty) - A_{T_n^\epsilon}(\kappa_a) \leq (F_b(\kappa_b + \epsilon) - F_b(\kappa_b) + \epsilon)T_n^\epsilon$.

We now pick a sequence of $\epsilon_n \rightarrow 0$, and take the diagonal subsequence of times:

$$T_1 = T_1^{\epsilon_1}, \quad T_{k+1} = \min_x \{T_x^{\epsilon_{k+1}} : T_x^{\epsilon_{k+1}} > T_k\}.$$

In what follows, we will be analyzing ordinary binned limit order books with finitely many bins. Define

$$\pi_t^b(k) = \frac{1}{t} \int_0^t \mathbf{1}\{\beta_u \in \text{bin } k\} du, \quad \pi_t^a(k) = \frac{1}{t} \int_0^t \mathbf{1}\{\alpha_u \in \text{bin } k\} du.$$

Here, we say that if $\beta_t = -\infty$ then β_t does not belong to any bin; similarly, if $\alpha_t = +\infty$ then α_t does not belong to any bin.

Let $\pi^b(\cdot)$ and $\pi^a(\cdot)$ be any limit point of $\pi_{T_n}^b(\cdot)$ and $\pi_{T_n}^a(\cdot)$, where T_n is the sequence identified in Lemma 6.1. Our goal will be to analyze π^b and π^a . We obtain the following characterization.

Lemma 6.2. *Let k_b, k_a be the bins containing κ_b and κ_a respectively. Let N be the total number of bins. Then π^b and π^a satisfy the following equations or inequalities:*

$$\sum_{k=1}^N \pi^b(k) \leq 1; \quad \sum_{k=1}^N \pi^a(k) \leq 1. \quad (6.1a)$$

$$\pi^a(k) = \pi^b(k) = 0, \quad k < k_b \text{ or } k > k_a. \quad (6.1b)$$

Equality holds in (6.1a) if $\kappa_b > -\infty$ and $\kappa_a < +\infty$.

Let $a(k)$, $b(k)$ be the probability that an arriving order falls into bin k . That is, $a(k) = \int_{bin\ k} f_a(p)dp$, and similarly for bids. Then

$$\begin{aligned} (1 - \sum_{l \leq k} \pi^a(l))b(k) - \pi^b(k)(\sum_{l \leq k} a(l)) = \\ \begin{cases} F_b(\kappa_b) - F_b(\inf\{x : x \in bin\ k\}), & k = k_b \\ 0, & k_b < k < k_a. \end{cases} \end{aligned} \quad (6.2a)$$

$$\begin{aligned} (1 - \sum_{l \geq k} \pi^b(l))a(k) - \pi^a(k)(\sum_{l \geq k} b(l)) = \\ \begin{cases} F_a(\sup\{x : x \in bin\ k\}) - F_a(\kappa_a), & k = k_a \\ 0, & k_b < k < k_a. \end{cases} \end{aligned} \quad (6.2b)$$

Proof. Equations (6.1) follow from the definition of κ_a and κ_b .

Equations (6.2) express the equation

$$\text{order arrivals} - \text{order departures} = \text{unfulfilled orders}.$$

The limiting number of unfulfilled orders in a given bin is given by Lemma 6.1. Let us show that the left-hand side represents the number of arrivals minus the number of departures.

Bids arrive into bin k if the lowest ask is in some bin $a_t > k$, and then they arrive at rate $b(k)$. Formally, whenever the bin a_t containing α_t satisfies $a_t > k$, the conditional probability that the next bid arrival will be into bin k is $b(k)$; if $a_t \leq k$, the conditional probability is 0. Thus, conditional on the amount of time that $a_t > k$, the number of bid arrivals into bin k is binomial with success probability $b(k)$. Since binomial random variables concentrate on their mean, we obtain the law of large numbers scaling above for the bid arrivals.

Similarly, bids depart from bin k if β_t is in bin k , and an ask arrives into some bin $l \leq k$. Therefore, conditional on the number of times that β_t is in bin k , the number of bid departures is binomial with success probability $\sum_{l \leq k} a(l)$.

The above is almost enough to determine the distributions π^b and π^a , except for the inequality in (6.1a). Our next goal will be to find some sufficient conditions to conclude $\kappa_b > -\infty$ and $\kappa_a < \infty$.

Lemma 6.3. *Suppose that the binning and arrival price distributions for the limit order book are such that there exist prices $x < y$ with the following properties:*

1. $0 < F_b(x) < F_b(y) < 1$, $0 < F_a(x) < F_a(y) < 1$;
2. The bin partition refines the partition $(-\infty, x] \cup (x, y] \cup (y, \infty)$;
3. $F_b(x) = 1 - F_a(y)$ and $F_b(y) = 1 - F_a(x)$.

Then κ_b and κ_a satisfy

$$1 - F_a(\kappa_a) = F_b(\kappa_b) \geq \frac{2X(1 - X) - (1 - Y)}{(1 - X) + (Y - X)},$$

where $X = F_b(x)$ and $Y = F_b(y)$.

In particular, if we may choose $Y = 1 - X$, then $F_b(\kappa_b) > 0$ and $F_a(\kappa_a) < \infty$.

Proof. We apply Lemma 6.2 to the 3-bin partition appearing in the statement. By Lemma 4.2, this provides a lower bound on κ_b , and an upper bound on κ_a , for the original problem. The final condition implies that in the 3-bin partition, the probability of a bid arrival in bin k is equal to X , $Y - X$, and $1 - Y$ for the three bins, and the probability of ask arrival in bin k is equal to $1 - Y$, $Y - X$, and X .

There will be six equations in (6.2). Note that in this scenario, $k_b = 4 - k_a$, and they cannot both be equal to 2 since α_t and β_t cannot be in the same bin.

Let us add together the following three pairs of equations: (6.2a) for bin 1 and (6.2b) for bin 3, (6.2a) for bin 2 and (6.2b) for bin 2, and finally (6.2a) for bin 3 and (6.2b) for bin 1. We obtain

$$\begin{pmatrix} X & 0 & 1 - Y \\ Y - X & Y - X + 1 - X & 0 \\ 1 - Y + 1 & 1 - Y & 1 - Y \end{pmatrix} \begin{pmatrix} \pi^a(1) + \pi^b(3) \\ \pi^a(2) + \pi^b(2) \\ \pi^a(3) + \pi^b(1) \end{pmatrix} = 2 \begin{pmatrix} X - F_b(\kappa_b) \\ Y - X \\ 1 - Y \end{pmatrix}.$$

Premultiplying by $(1 - X + Y - X, 1 - Y, (1 - Y) - 2X(1 - X))$ and observing that $\sum \pi^a(k) + \pi^b(k) \leq 2$, we obtain

$$F_b(\kappa_b) \geq \frac{2X(1 - X) - (1 - Y)}{(1 - X) + (Y - X)},$$

as required. The final assertion follows because for $Y = 1 - X$ with $X < 1/2 < Y$, the numerator is strictly positive. (Note the denominator is always positive.)

We now have obtained sufficient conditions for Lemma 6.2 to provide us with the full description of the (discrete) distributions π^a and π^b , *if we knew the value of κ_b .*

7. Proof of Theorem 3.2

Let us summarize what we know so far. We are interested in finding the value of κ_b in an ordinary (unbinned) limit order book. We know (Lemma 5.2) that it can be obtained by considering limit order books with smaller and smaller bins. Whenever we have a finite number of bins, Lemma 6.2 tells us how to find π^b and π^a , a pair of distributions supported on the bins between κ_b and κ_a . We expect π^b and π^a to be the steady-state distribution of the (bin containing the) rightmost bid and of the (bin containing the) leftmost ask respectively, although we have only shown that it is the limiting distribution along a special sequence of times.

We now observe the following. Suppose we knew the values of κ_b and κ_a . Then Lemma 6.2 would allow us to compute π^b and π^a , and we could rediscover κ_b and κ_a as the boundaries of their support. We will now see that the requirements that (i) Lemma 6.2 should hold and (ii) (κ_b, κ_a) should be the support of the resulting distributions are enough to determine κ_b and κ_a .

Proof of Theorem 3.2. Reparametrize coordinates so that all arrivals happen on $[0, 1]$.

We consider a sequence of ordinary binned limit order books, \mathcal{L}^N , which differ in the binning partitions Π^N that they use. We require that Π^N refines Π^{N-1} , and that each bin of Π^N has width $\leq \frac{1}{N}$, and also $a(k), b(k) \leq \frac{1}{N}$. Here, $a(k)$ is the probability that an arriving ask enters bin k , and similarly for $b(k)$ and bids.

Suppose that in such a limit order book we do not know κ_b and κ_a , but do know the bins $k_b = k_b(N)$ and $k_a = k_a(N)$ into which they fall. Then (6.2) gives equations for $\pi^a(k)$ and $\pi^b(k)$ for all $k \neq k_a, k_b$, plus the following inequalities:

1. $0 \leq \pi^b(k) \leq b(k)$, $0 \leq \pi^a(k) \leq a(k)$ for any k ;
2. $0 \leq \pi^b(k_a) \leq \frac{\pi^a(k_a)}{F_a(\kappa_a)} b(k_a)$, $0 \leq \pi^a(k_b) \leq \frac{\pi^b(k_b)}{1 - F_b(\kappa_b)} a(k_b)$.

Now consider taking $N \rightarrow \infty$. By (1), $\pi^b(k_b)$ and $\pi^a(k_a) \rightarrow 0$. Then by (2),

$$\frac{\pi^b(k_a)}{b(k_a)} \rightarrow 0, \quad \frac{\pi^a(k_b)}{a(k_b)} \rightarrow 0. \quad (7.1)$$

Note that these are the Radon-Nikodym derivatives of π^b with respect to the bid arrival distribution f_b (discretized to bins), and of π^a with respect to f_a .

Consider now $\pi^b(k_b + 1)$. For it we have

$$\pi^b(k_b + 1) \left(\sum_{k \leq k_b + 1} a(k) \right) = \left(\sum_{k > k_b + 1} \pi^a(k) \right) b(k_b + 1).$$

By considerations similar to (1), $\pi^a(k) \rightarrow 0$ uniformly as $N \rightarrow \infty$ for any single bin k . Consequently, as $N \rightarrow \infty$, $\sum_{k > k_b(N)+1} \pi^a(k) = 1 - \pi^a(k_b(N)) - \pi^a(k_b(N) + 1) \rightarrow 1$, and we obtain

$$\frac{\pi^b(k_b(N) + 1)}{b(k_b(N) + 1)} \rightarrow \left(\sum_{k \leq k_b + 1} a(k) \right)^{-1} \quad (7.2)$$

and similarly for asks.

An identical calculation yields $\frac{\pi^b(k_b(N)+2)}{b(k_b(N)+2)} \rightarrow \left(\sum_{k \leq k_b+2} a(k) \right)^{-1}$, from which

$$\frac{\pi^b(k_b(N) + 1)}{b(k_b(N) + 1)} - \frac{\pi^b(k_b(N) + 2)}{b(k_b(N) + 2)} \rightarrow \frac{a(k_b(N) + 2)}{\left(\sum_{k \leq k_b(N)+1} a(k) \right) \left(\sum_{k \leq k_b(N)+2} a(k) \right)}, \quad (7.3)$$

and similarly for asks. Note that we could have the summation running to $k_b(N)$ in the denominator, since $a(k) \rightarrow 0$ uniformly for any single bin k .

As $N \rightarrow \infty$, the derivatives $\frac{\pi^b(k)}{b(k)}$, $k = 1, \dots, N$ and $\frac{\pi^a(k)}{a(k)}$, $k = 1, \dots, N$ are bounded between 0 and $\frac{1}{F_a(\kappa_b)}$, resp. $\frac{1}{F_b(\kappa_a)}$, and hence converge along some subsequence. Consider any such pair of subsequential limits (we will shortly see that it is unique), ϖ^b and ϖ^a . We find

$$0 \leq \varpi^b(x), \varpi^a(x);$$

$$\varpi^b(x) = \varpi^a(x) = 0, \quad x < \kappa_b \text{ or } x > \kappa_a;$$

$$\int_{\kappa_b}^{\kappa_a} \varpi^b(x) f_b(x) dx = \int_{\kappa_b}^{\kappa_a} \varpi^a(x) f_a(x) dx = 1;$$

$$\varpi^b(x) F_a(x) = \left(\int_x^{\kappa_a} \varpi^a(y) f_a(y) dy \right), \quad \kappa_b < x < \kappa_a$$

$$\varpi^a(x) (1 - F_b(x)) = \left(\int_{\kappa_b}^x \varpi^b(y) f_b(y) dy \right), \quad \kappa_b < x < \kappa_a;$$

$$\begin{aligned}\varpi^b(\kappa_b) &= \frac{1}{F_a(\kappa_b)}, & \frac{d}{dx}\varpi^b(x)|_{x=\kappa_b} &= -\frac{f_a(\kappa_b)}{F_a(\kappa_b)^2}, \\ \varpi^a(\kappa_a) &= \frac{1}{1-F_b(\kappa_a)}, & \frac{d}{dx}\varpi^a(x)|_{x=\kappa_a} &= \frac{f_b(\kappa_a)}{(1-F_b(\kappa_a))^2}, \\ \varpi^b(x) &\rightarrow 0, \quad x \uparrow \kappa_a; & \varpi^a(x) &\rightarrow 0, \quad x \downarrow \kappa_b.\end{aligned}$$

It remains to observe that the pair of integral equations can be converted into a pair of differential equations for $\varpi^b(x)$. Indeed,

$$(F_a(x)\varpi^b(x))' = -\varpi^a(x)f_a(x) = -\frac{f_a(x)}{1-F_b(x)} \int_{\kappa_b}^x \varpi^b(y)f_b(y)dy,$$

and hence

$$\left(-\frac{f_a(x)}{1-F_b(x)} (F_a(x)\varpi^b(x))'\right)' = \varpi^b(x)f_b(x).$$

For this second-order ODE, we have two initial conditions – the values $\varpi^b(\kappa_b)$ and $\frac{d}{dx}\varpi^b(x)|_{x=\kappa_b}$. This allows us to find a solution for the ODE given κ_b , for each value of κ_b . Recall, however, that we have an additional constraint $\varpi^b(x) \rightarrow 0$ as $x \rightarrow \kappa_a$, where κ_a and κ_b are related via $F_b(\kappa_b) = 1 - F_a(\kappa_a)$. It is not difficult to see that there can only be one value of κ_b that is consistent with the ODE and the additional constraint of ϖ vanishing at $\kappa_a = F_a^{-1}(1 - F_b(\kappa_b))$.

7.1. Calculations for uniform distribution

For the case of the uniform distribution, $f_a = f_b = \mathbf{1}_{[0,1]}$, we can take the calculations a step further by solving the above differential equation. We obtain

$$\left(-\frac{1}{1-x} (x\varpi^b(x))'\right)' = \varpi^b(x), \quad \varpi^b(\kappa_b) = \frac{1}{\kappa_b}, \quad \frac{d}{dx}\varpi^b(x)|_{x=\kappa_b} = -\frac{1}{\kappa_b^2},$$

which can be solved explicitly to give

$$\varpi^b(x) = (1-\kappa) \left(\frac{1}{x} + \log \left(\frac{1-x}{x} \right) \right), \quad x \in (\kappa, 1-\kappa).$$

The value of $\kappa = \kappa_b$ is given as follows. Let w be the unique solution of $we^w = e^{-1}$; then $\kappa = \frac{w}{w+1} \approx 0.217$.

Figure 1 compares the empirical distribution of the location of β_t when we consider 100 bins, and the curve ϖ^b given above. The close agreement between the two curves supports Conjecture 1. Further support is given by Figure 2, which shows the total number of bids in bins to the right of the threshold as a function of time. The plot of the maximal value observed up to time t as a function of t is seen not to grow linearly, supporting the conjecture.

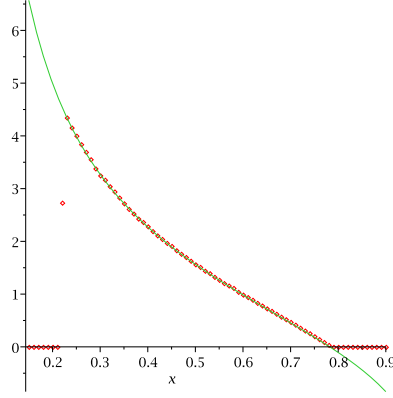


FIGURE 1: Empirical distribution of the location of β_t with uniform bid and ask arrivals and 100 bins, and the predicted density for the unbinned model.

8. Recurrence

In this section, our goal is to prove results similar to Conjecture 1. While we will not be able to derive recurrence when there is an infinite supply of bids and asks at κ_b and κ_a , we will be able to derive it for smaller subintervals.

Theorem 8.1. *Let the arrivals of both bids and asks are Poisson of rate 1 in time, with densities f_b and f_a respectively in price. Suppose that there exist values x and y such that $F_b(y) < F_b(x) + F_a(x)$ and $F_a(y) < F_a(x) + (1 - F_b(y))$. (For example, we may have $F_a(x) = F_b(x) > 1/3$ and $F_a(y) = F_b(y) < 2/3$.) Let \mathcal{L} be a limit order book whose initial state is such that there are infinitely many bid orders at x and infinitely many ask orders at y . Letting the state of \mathcal{L} be described by the bids and asks in (x, y) , it is a positive Harris recurrent Markov chain.*

Proof. Consider the bids in (x, y) . They arrive at rate at most $b(3) = F_b(y) - F_b(x)$. Moreover, whenever there are any bids in bin 3, they depart at rate at least $F_a(x)$. Consequently, the number of bids in bin 3 is (stochastically) bounded above by a geometric random variable with parameter $\frac{F_b(y) - F_b(x)}{F_a(x)} < 1$. Similarly, the number of asks in bin 3 is bounded above by a geometric random variable with parameter $\frac{F_a(y) - F_a(x)}{1 - F_b(y)} < 1$. This suffices to prove the claim.

Note that in particular this result gives an upper bound on the value of κ_b (and a lower bound on κ_a).

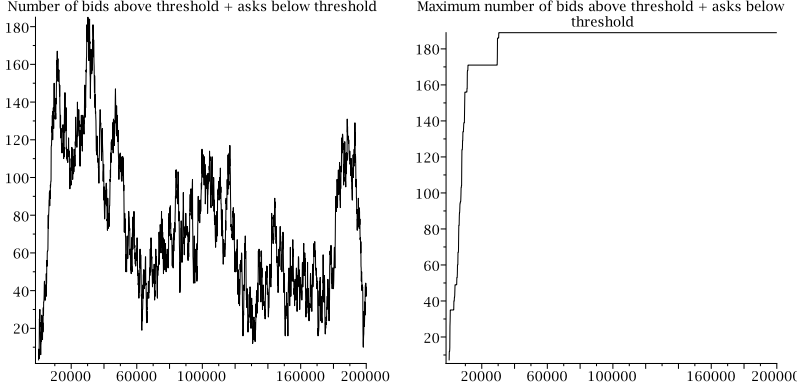


FIGURE 2: Left: total number of bids above bin $k_b = 21$ and asks below bin $k_a = 79$, as a function of time. Right: running maximum of this quantity, as a function of time. The last jump of the running maximum occurs at arrival 30,309 of 200,000.

We can prove a slightly stronger result for binned limit order books.

Theorem 8.2. *Suppose arriving orders are equally likely to be bids and asks, and $f_a = f_b = \mathbf{1}_{[0,1]}$. Let $\epsilon > 0$, and consider the binned limit order book with 5 bins of sizes $1/5 + \epsilon$, $1/5 - \epsilon$, $1/5$, $1/5 - \epsilon$, $1/5 + \epsilon$ whose initial state has infinitely many bids in bin 1 and infinitely many asks in bin 5. This limit order book (considered on the middle three bins only) is positive recurrent.*

Proof. Let $X_t \in \mathbb{Z}^3$ be the Markov chain describing their state. We let $|X_t(i)|$ denote the number of orders in bin i , and its sign correspond to bids (+) or asks (-).

The evolution of the system depends on the bins containing the rightmost bid and of the leftmost ask; call these b_t and a_t respectively. There are 10 possible combinations, which we denote $+++$, $++-$, $+--$, $---$, $++0$, $+0-$, $0--$, $+00$, $00-$, and 000 . The signs should be thought of as the signs of $X_t(i)$, although we do not distinguish e.g. $0++$ from $+++$. Note that 000 corresponds to the middle three bins being empty.

Consider the (vector) drift of X , that is, $\mathbb{E}[X_{t+1} - X_t | X_t]$. As we mentioned, this depends only on the region F to which X_t belongs, where F comes from the list of possible descriptions for the pair $(b(t), a(t))$. We will not be interested in the drift

when $F = 000$. The drifts are as follows:

$$\begin{aligned}
\Delta_{+++} &= (1/5 - \epsilon, 1/5, -(4/5 - \epsilon)), & \Delta_{---} &= (4/5 - \epsilon, -1/5, -(1/5 - \epsilon)), \\
\Delta_{++-} &= (1/5 - \epsilon, -3/5, 2/5 - \epsilon), & \Delta_{+--} &= (-(2/5 - \epsilon), 3/5, -(1/5 - \epsilon)), \\
\Delta_{++0} &= (1/5 - \epsilon, -3/5, 0), & \Delta_{0--} &= (0, 3/5, -(1/5 - \epsilon)), \\
\Delta_{+0-} &= (-(2/5 - \epsilon), 0, 2/5 - \epsilon), & \Delta_{+00} &= (-(2/5 - \epsilon), 0, 0), \\
\Delta_{00-} &= (0, 0, 2/5 - \epsilon).
\end{aligned}$$

We will now show that X is positive recurrent by constructing a Lyapunov function for it. Note that the jumps of X are bounded by 1. We define $\mathcal{L}(x)$ by

$$\mathcal{L}(x) = \min(\langle x, v_F \rangle), F \in \{++-, +--, ++0, +00, +0-, 0--, 00-\}.$$

We will specify the vectors v_F shortly. The set of possible values of F is the set of relative positions of bid and ask above, except for 000 (the origin).

The level sets of this Lyapunov function are polyhedra with outer normals v_F . The vectors v_F will be picked so that a point x on the face of the level polyhedron with outer normal v_F belongs either to the orthant F , or (if F is not an orthant, i.e. if it contains 0) to one of the orthants adjacent to F . In this way, it will be sufficient to construct v_F so that $\langle \Delta_{\tilde{F}}, v_F \rangle < 0$ whenever \tilde{F} agrees with F at all the nonzero places of F . By compactness of the level sets, this will guarantee that

$$\exists K > 0 \text{ such that } \mathbb{E}[\mathcal{L}(X_{t+1}) - \mathcal{L}(X_t) | X_t] < -\epsilon < 0 \text{ whenever } \mathcal{L}(X_t) > K.$$

Together with the fact that the jumps of \mathcal{L} are clearly bounded above (because the jumps of X are, and \mathcal{L} is Lipschitz), this gives the Foster-Lyapunov criterion for positive recurrence as used e.g. in [1, Proposition 4.4].

It can be checked that the choice

$$\begin{aligned}
v_{+++} &\equiv (1, 1, 1), & v_{++-} &\equiv (1, 1, -1), \\
v_{+--} &\equiv (1, -1, -1), & v_{---} &\equiv (-1, -1, -1), \\
v_{++0} = v_{+00} &\equiv (4/3, 1, 2/3), & v_{+0-} &\equiv (1, -4/5, -9/5), \\
v_{0--} = v_{00-} &\equiv (-2, -3, -4).
\end{aligned}$$

satisfies all of these constraints. Figure 3 shows the corresponding level set of \mathcal{L} .

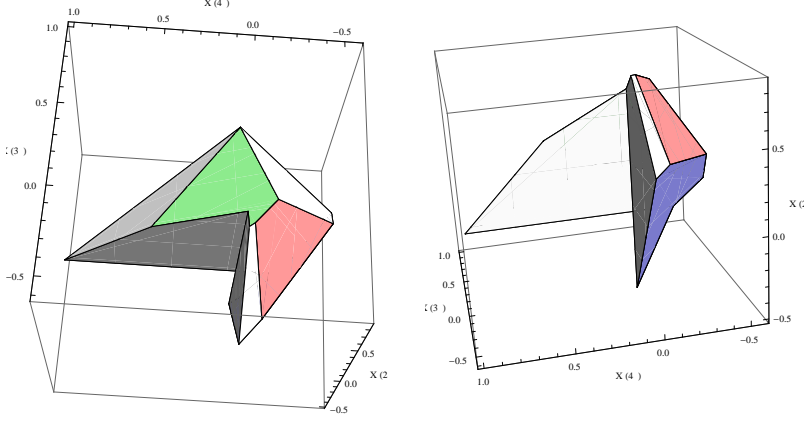


FIGURE 3: Two views of the level set $P \equiv \{x : \mathcal{L}(x) = 1\}$. P has 15 vertices $(0,0,0)$, $(0,1,0)$, $(0,0,1)$, $(\frac{1}{2}, 0, \frac{1}{2})$, $(\frac{45}{58}, \frac{2}{29}, -\frac{9}{58})$, $(\frac{6}{7}, -\frac{1}{7}, 0)$, $(\frac{29}{34}, -\frac{2}{17}, -\frac{1}{34})$, $(\frac{3}{4}, 0, 0)$, $(\frac{11}{50}, \frac{6}{25}, -\frac{27}{50})$, $(0, \frac{3}{7}, -\frac{4}{7})$, $(\frac{11}{26}, -\frac{6}{13}, -\frac{3}{26})$, $(\frac{2}{5}, -\frac{3}{5}, 0)$, $(0, -\frac{1}{3}, 0)$, $(-\frac{1}{2}, 0, 0)$, $(0, 0, -\frac{1}{4})$, and 10 faces (defined as ordered sets of vertices, possibly not all with the same orientation) $\{4, 3, 2\}$, $\{5, 2, 10, 9\}$, $\{7, 6, 12, 11\}$, $\{1, 3, 4, 8\}$, $\{1, 8, 6, 12, 14\}$, $\{1, 3, 2, 10, 15\}$, $\{1, 15, 14\}$, $\{7, 5, 9, 11\}$, $\{2, 4, 8, 6, 7, 5\}$, $\{9, 10, 15, 14, 12, 11\}$.

This result does not directly translate into a statement about recurrence of an unbinned limit order book. However, it can be used to derive some bounds on κ_b and κ_a for an ordinary (unbinned) limit order book with uniform arrivals of bids and asks. Analysis similar to that of Lemma 5.2 gives, by looking at strict limit order books with 4 bins, $\kappa_b < 1/4$ and $\kappa_a > 3/4$. We do not go into details here, since for this case we have already computed the value of κ_b precisely in Section 7.1.

One further indication that the positive Harris recurrence holds is obtained by plotting the empirical density of the joint location of the highest bid and the lowest ask. Figure 4 presents the plots obtained by simulation. The plots suggest that there is a limiting surface describing the joint density, although we have been unable to obtain an expression for it.

9. Discussion and future work

We have presented a model of a limit order book applicable on relatively short time scales, during which the price is relatively stationary, order arrivals can be

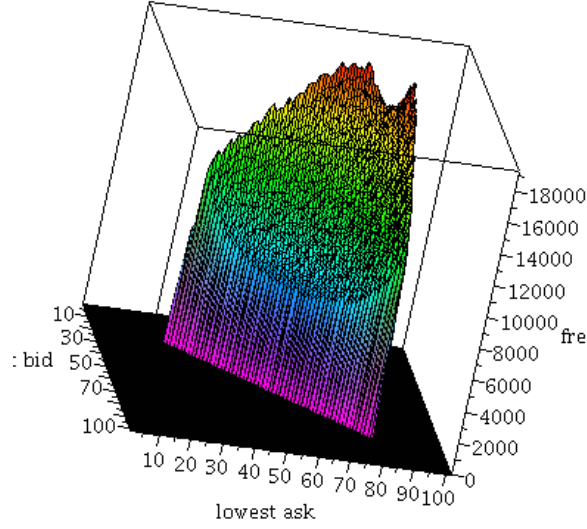


FIGURE 4: Joint distribution of the rightmost bid and the leftmost ask.

modeled as having distributions independent of the state of the limit order book, and abandonments can be ignored (either as background noise or as orders that are never executed). We have obtained a phase transition, which suggests that at such a time orders in a narrow band around the price get executed (more and more rarely as we move away from the price), while orders farther away from the price will never be reached. We have determined this band in terms of the distribution of the prices of arriving orders. Modulo a conjectural result on recurrence, we have also determined how frequently orders at a given distance away from the optimal price get executed.

There are two major directions in which this work could be extended.

First, it would be desirable to understand finer features of the model. In particular, we would like to gain an understanding of the joint distribution of the highest bid and lowest ask, which would also allow us to understand the distribution of the bid-ask spread in this model. It would also be interesting to examine the limiting shape of the book. Note that the distribution of the rightmost bid gives an indication of the expected number of bids and asks at any given price, but we would like to understand the following: conditional on β_t being in bin b_t and α_t being in bin a_t , what is the number of bids in bin $b_t - k$? What is the number of asks in bin $a_t + k$? Can we

reproduce some of the empirical results concerning these shapes?

The left-hand side of Figure 5, we present the number of bids in the bin containing the highest bid, as a function of time. The occasional spikes are not surprising, because we expect the highest bid to occasionally be in the threshold bin, which contains large numbers of orders. The right-hand side presents the average number of bids at and near the bid price, *when the bid price is high*. We see that there often are no bid orders near the highest waiting bid, when the arrival distribution is uniform. This may, however, change if we consider different arrival distributions.

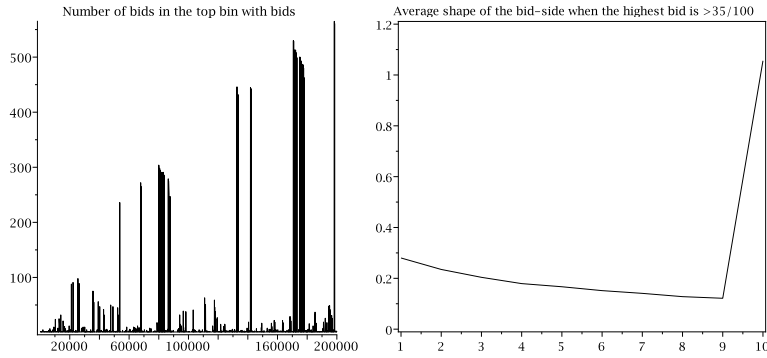


FIGURE 5: The number of bids in the bin with the highest bid; the average shape of the bid side of the limit order book when the bid price is high.

Second, it would be desirable to change the model so that it would more closely agree with actual limit order books. In particular, we would like to be able to incorporate order abandonment, non-unit-sized orders, and order distributions that depend on the price. Order abandonment fundamentally changes our analysis by removing the hard phase transition and getting rid of monotonicity properties; it is likely that it requires a different formulation of the model. However, different arrival distributions may be amenable to analysis once we have a better understanding of the shape of the limit order book around its price.

Acknowledgements

The author would like to thank Yuri Suhov and Frank Kelly for introducing her to limit order books and for many insightful conversations, and Vlada Limic and Florian

Simatos for interesting conversations about this and related models. Special thanks to Daniel Whalen for help generating the images in Figure 3.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship and Award No. 1204311.

References

- [1] BRAMSON, M. (2006). *Stability and Heavy Traffic Limits for Queueing Networks: St. Flour Lectures Notes*. Springer. <http://www.math.duke.edu/~rtd/CPSS2007/Bramson.pdf>.
- [2] CONT, R. AND DE LARRARD, A. Price dynamics in a markovian limit order market. Working paper 2010. <http://ssrn.com/abstract=1735338>.
- [3] CONT, R., STOIKOV, S. AND TALREJA, R. (2010). A stochastic model for order book dynamics. *Operations Research* **58**, 549–563.
- [4] GODE, D. K. AND SUNDER, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy* **101**, 119–137.
- [5] GOULD, M. D., PORTER, M. A., WILLIAMS, S., McDONALD, M., FENN, D. J. AND HOWISON, S. D. Limit order books. In preparation 2012. [arXiv:1012.0349v3](https://arxiv.org/abs/1012.0349v3).
- [6] SIMATOS, F. AND AIDEKON, E. Coupling between a stochastic model of a limit order book and branching random walks. Submitted 2011. [arXiv:1210.7062](https://arxiv.org/abs/1210.7062).